

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 10/528,948
Source: PCT
Date Processed by STIC: 08/03/2005

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 08/03/2005

PATENT APPLICATION: US/10/528,948

TIME: 13:57:54

Input Set : N:\Cr3\RULE60\10528948.raw

Output Set: N:\CRF4\08032005\J528948.raw

CD

```

1 <110> APPLICANT: THE GENERAL HOSPITAL CORPORATION
2   SHELLEY, CARL SIMON
3   FAROKHZAD, OMID C.
4 <120> TITLE OF INVENTION: METHODS FOR DIAGNOSING AND TREATING TUMORS AND SUPPRESSING
5   PROMOTERS
6 <130> FILE REFERENCE: M00765.70064
7 <140> CURRENT APPLICATION NUMBER: 10/528,948
8 <141> CURRENT FILING DATE: 2005-03-23
9 <150> PRIOR APPLICATION NUMBER: PCT/US03/30213
10 <151> PRIOR FILING DATE: 2003-09-23
11 <150> PRIOR APPLICATION NUMBER: US 60/412,964
12 <151> PRIOR FILING DATE: 2002-09-23
13 <160> NUMBER OF SEQ ID NOS: 28
14 <170> SOFTWARE: PatentIn version 3.2
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 1879
18 <212> TYPE: DNA
19 <213> ORGANISM: Homo sapiens sialophorin
20 <400> SEQUENCE: 1
21   gcctcgggag gtggtggagt gacctggccc cagtgtctgcg tccttatcag ccgagccggt      60
22   cccagctctt gctcctgcct gtttgccctg aaatggccac gcttctcctt ctccttgggg      120
23   tgctgggtgt aagcccagac gctctgggga gcacaacagc agtgcagaca cccacctccg      180
24   gagagccttt ggtctctact agcgagcccc tgagctcaaa gatgtacacc acttcaataa      240
25   caagtgaacc taaggccgac agcactgggg accagacctc agccctacct cctcaactt      300
26   ccatcaatga gggatcccc ctcttgactt ccattggtgc cagcactggt tcccctttac      360
27   ctgagccaac aacctaccag gaagtttcca tcaagatgtc atcagtgcc caggaaaccc      420
28   ctcatgcaac cagtcctcct gctgttccca taacagcaaa ctctctagga tcccacaccg      480
29   tgacaggtgg aaccataaca acgaactctc cagaaacctc cagtaggacc agtggagccc      540
30   ctgttaccac ggcagctagc tctctggaga cctccagagg cacctctgga ccccctctta      600
31   ccatggcaac tgtctctctg gagacttcca aaggcacctc tggaccccct gttaccatgg      660
32   caactgactc tctggagacc tccactggga ccaactggacc ccctgttacc atgacaactg      720
33   gctctctgga gccctccagc ggggccagtg gaccccaggt ctctagcgta aaactatcta      780
34   caatgatgtc tccaacgacc tccaccaacg caagcactgt gcccttccgg aaccagatg      840
35   agaactcacg aggcattgct ccagtggtctg tgcttgtggc cctgctggcg gtcatagtcc      900
36   tcgtggctct gctcctgctg tggcgccggc ggcagaagcg gcggactggg gccctcgtgc      960
37   tgagcagagg tggcaagcgt aacgggggtg tggacgcctg ggctgggcca gccacaggtcc      1020
38   ctgaggaggg ggccgtgaca gtgaccgtgg gaggggtccg gggcgacaag ggctctgggt      1080
39   tccccgatgg ggaggggtct agccgtcggc ccacgctcac cactttcttt ggcagacgga      1140
40   agtctcgcca gggctccctg gcgatggagg agctgaagtc tgggtcaggc cccagcctca      1200
41   aaggggagga ggagccactg gtggccagtg aggatggggc tgtggacgcc ccagctcctg      1260
42   atgagcccga agggggagac ggggctgccc cttaagtgtc ggtgaatagt gaggctggag      1320
43   gccggaatct cagccagcct ccagcacctt ccctctcacc atcccactgc cccctcgctc      1380
44   ccatgtttcc acccggcacc ctgatcctca cccgaatctc cttttttttt ttcttttgag      1440

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/528,948

DATE: 08/03/2005

TIME: 13:57:54

Input Set : N:\CrF3\RULE60\10528948.raw

Output Set: N:\CRF4\08032005\J528948.raw

```

45      acagagtttc gctttgtcgc ccaggctgga gtgcaatgca cgatctcagt tcaactgcaac      1500
46      ctctgcctcc taagttcagg cgattctcct gcctcagctt cccgagtaac tgagattaca      1560
47      ggcacccacc accatgcccc gctgcttttt tgtatttttg gtagagatgg ggtttcacca      1620
48      tgttggctag gctgggtctca aactcctgac ctcaggtgat ctacctgcct cagcctccca      1680
49      aagtgtctgag attacagaca tgagcctccg cgccttgccct cctcaccacac ctcttcactc      1740
50      tgaatcctca tgaggcttct cagccctgga tttcctgctg ccatacctcac ccagcaccca      1800
51      caactagcgc ctgggcaggg cagggctggc acctctcaac gtctgtggac tgaatgaata      1860
52      aaccctcctc atccacccc                                1879
54 <210> SEQ ID NO: 2
55 <211> LENGTH: 400
56 <212> TYPE: PRT
57 <213> ORGANISM: Homo sapiens sialophorin
58 <400> SEQUENCE: 2
59      Met Ala Thr Leu Leu Leu Leu Gly Val Leu Val Val Ser Pro Asp
60      1          5          10          15
61      Ala Leu Gly Ser Thr Thr Ala Val Gln Thr Pro Thr Ser Gly Glu Pro
62      20          25          30
63      Leu Val Ser Thr Ser Glu Pro Leu Ser Ser Lys Met Tyr Thr Thr Ser
64      35          40          45
65      Ile Thr Ser Asp Pro Lys Ala Asp Ser Thr Gly Asp Gln Thr Ser Ala
66      50          55          60
67      Leu Pro Pro Ser Thr Ser Ile Asn Glu Gly Ser Pro Leu Trp Thr Ser
68      65          70          75          80
69      Ile Gly Ala Ser Thr Gly Ser Pro Leu Pro Glu Pro Thr Thr Tyr Gln
70      85          90          95
71      Glu Val Ser Ile Lys Met Ser Ser Val Pro Gln Glu Thr Pro His Ala
72      100         105         110
73      Thr Ser His Pro Ala Val Pro Ile Thr Ala Asn Ser Leu Gly Ser His
74      115         120         125
75      Thr Val Thr Gly Gly Thr Ile Thr Thr Asn Ser Pro Glu Thr Ser Ser
76      130         135         140
77      Arg Thr Ser Gly Ala Pro Val Thr Thr Ala Ala Ser Ser Leu Glu Thr
78      145         150         155         160
79      Ser Arg Gly Thr Ser Gly Pro Pro Leu Thr Met Ala Thr Val Ser Leu
80      165         170         175
81      Glu Thr Ser Lys Gly Thr Ser Gly Pro Pro Val Thr Met Ala Thr Asp
82      180         185         190
83      Ser Leu Glu Thr Ser Thr Gly Thr Thr Gly Pro Pro Val Thr Met Thr
84      195         200         205
85      Thr Gly Ser Leu Glu Pro Ser Ser Gly Ala Ser Gly Pro Gln Val Ser
86      210         215         220
87      Ser Val Lys Leu Ser Thr Met Met Ser Pro Thr Thr Ser Thr Asn Ala
88      225         230         235         240
89      Ser Thr Val Pro Phe Arg Asn Pro Asp Glu Asn Ser Arg Gly Met Leu
90      245         250         255
91      Pro Val Ala Val Leu Val Ala Leu Leu Ala Val Ile Val Leu Val Ala
92      260         265         270
93      Leu Leu Leu Leu Trp Arg Arg Arg Gln Lys Arg Arg Thr Gly Ala Leu
94      275         280         285

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/528,948

DATE: 08/03/2005

TIME: 13:57:54

Input Set : N:\Crif3\RULE60\10528948.raw

Output Set: N:\CRF4\08032005\J528948.raw

```

95      Val Leu Ser Arg Gly Gly Lys Arg Asn Gly Val Val Asp Ala Trp Ala
96      290                                295                                300
97      Gly Pro Ala Gln Val Pro Glu Glu Gly Ala Val Thr Val Thr Val Gly
98      305                                310                                315                                320
99      Gly Ser Gly Gly Asp Lys Gly Ser Gly Phe Pro Asp Gly Glu Gly Ser
100      325                                330                                335
101      Ser Arg Arg Pro Thr Leu Thr Thr Phe Phe Gly Arg Arg Lys Ser Arg
102      340                                345                                350
103      Gln Gly Ser Leu Ala Met Glu Glu Leu Lys Ser Gly Ser Gly Pro Ser
104      355                                360                                365
105      Leu Lys Gly Glu Glu Glu Pro Leu Val Ala Ser Glu Asp Gly Ala Val
106      370                                375                                380
107      Asp Ala Pro Ala Pro Asp Glu Pro Glu Gly Gly Asp Gly Ala Ala Pro
108      385                                390                                395                                400

```

110 <210> SEQ ID NO: 3

111 <211> LENGTH: 1893

112 <212> TYPE: DNA

113 <213> ORGANISM: Homo sapiens

114 <400> SEQUENCE: 3

```

115      cttccctgcc tccctcaggt cccagctctt gctcctgcct gtttgccctgg aaatggccac      60
116      gcttctcctt ctccctgggg tgctggtggt aagcccagac gctctgggga gcacaacagc      120
117      agtgcagaca cccacctcgg gagagccttt ggtctctact agcgagcccc tgagctcaaa      180
118      gatgtacacc acttcaataa caagtgaccc taaggccgac agcactgggg accagacctc      240
119      agccctacct cctcaactt ccatcaatga gggatccctt ctttggactt ccattggtgc      300
120      cagcactggg tccccctttac ctgagccaac aacctaccag gaagtttcca tcaagatgtc      360
121      atcagtcccc caggaaaccc ctcatgcaac cagtcacctt gctgttccca taacagcaaa      420
122      ctctctagga tcccacaccg tgacaggtgg aaccataaca acgaactctc cagaaacctc      480
123      cagtaggacc agtggagccc ctgttaccac ggcagctagc tctctggaga cctccagagg      540
124      cacctctgga cccctcttta ccatggcaac tgtctctctg gagacttcca aaggcacctc      600
125      tggacccctt gttaccatgg caactgactc tctggagacc tccactggga ccactggacc      660
126      cctgtttacc atgacaactg gctctctgga gccctccagc ggggcccagt gaccccaggt      720
127      ctctagcgta aaactatcta caatgatgtc tccaacgacc tccaccaacg caagcactgt      780
128      gcccttccgg aaccacagatg agaactcaag aggcattgctg ccagtggctg tgcttgtggc      840
129      cctgctggcg gtcatagtcc tcgtggtctt gctcctgctg tggcgccggc ggcagaagcg      900
130      gcggactggg gccctcgtgc tgagcagagg cggcaagcgt aacggggtgg tggacgcctg      960
131      ggctgggcca gcccaggtcc ctgaggaggg ggccgtgaca gtgaccgtgg gaggtccgg      1020
132      gggcgacaag ggctctgggt tccccgatgg ggaggggtct agccgtcggc ccacgctcac      1080
133      cactttcttt ggcagacgga agtctcgcca gggctccctg gcgatggagg agctgaagtc      1140
134      tgggtcaggc cccagcctca aaggggagga ggagccactg gtggccagtg aggatggggc      1200
135      tgtggacgcc ccagctcctg atgagccoga agggggagac ggggctgccc cttaagtgtc      1260
136      ggtgaatagt gaggtggag gccggaatct cagccagcct ccagcacctt ccctctcacc      1320
137      atcccactgc cccctcgtct ccatgtttcc acccggcacc ctgatcctca cccgaatctc      1380
138      cttttttttt ttcttttgag acagagtttc gctttgtcgc ccaggctgga gtgcaatgca      1440
139      cgatctcagt tcaactgcaac ctctgcctcc taagttcagg cgattctcct gcctcagctt      1500
140      cccgagtaac tgagattaca ggcacccacc accatgcccc gctgcttttt tgtatttttg      1560
141      gtagagatgg ggtttcacca tgttggctag gctggtctca aactcctgac ctcaggtgat      1620
142      ctacctgcct cagcctccca aagtgtctgag attacagaca tgagcctcgg cgccttgctt      1680
143      cctcaccac ctcttcactc tgaatctca tgaggcttct cagccctgga tttcctgctg      1740
144      ccatcctcac ccagacacca caactagcgc ctgggcaggg cagggtgagg acctctcaac      1800

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/528,948

DATE: 08/03/2005

TIME: 13:57:54

Input Set : N:\Crff3\RULE60\10528948.raw

Output Set: N:\CRF4\08032005\J528948.raw

```

145      gtctgtggac tgaatgaata aaccctcctc atccaccctt aaaaaaaaaa aaaaaaaaaa      1860
146      aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaa                                     1893
148 <210> SEQ ID NO: 4
149 <211> LENGTH: 400
150 <212> TYPE: PRT
151 <213> ORGANISM: Homo sapiens
152 <400> SEQUENCE: 4
153      Met Ala Thr Leu Leu Leu Leu Leu Gly Val Leu Val Val Ser Pro Asp
154      1              5              10              15
155      Ala Leu Gly Ser Thr Thr Ala Val Gln Thr Pro Thr Ser Gly Glu Pro
156      20              25              30
157      Leu Val Ser Thr Ser Glu Pro Leu Ser Ser Lys Met Tyr Thr Thr Ser
158      35              40              45
159      Ile Thr Ser Asp Pro Lys Ala Asp Ser Thr Gly Asp Gln Thr Ser Ala
160      50              55              60
161      Leu Pro Pro Ser Thr Ser Ile Asn Glu Gly Ser Pro Leu Trp Thr Ser
162      65              70              75              80
163      Ile Gly Ala Ser Thr Gly Ser Pro Leu Pro Glu Pro Thr Thr Tyr Gln
164      85              90              95
165      Glu Val Ser Ile Lys Met Ser Ser Val Pro Gln Glu Thr Pro His Ala
166      100             105             110
167      Thr Ser His Pro Ala Val Pro Ile Thr Ala Asn Ser Leu Gly Ser His
168      115             120             125
169      Thr Val Thr Gly Gly Thr Ile Thr Thr Asn Ser Pro Glu Thr Ser Ser
170      130             135             140
171      Arg Thr Ser Gly Ala Pro Val Thr Thr Ala Ala Ser Ser Leu Glu Thr
172      145             150             155             160
173      Ser Arg Gly Thr Ser Gly Pro Pro Leu Thr Met Ala Thr Val Ser Leu
174      165             170             175
175      Glu Thr Ser Lys Gly Thr Ser Gly Pro Pro Val Thr Met Ala Thr Asp
176      180             185             190
177      Ser Leu Glu Thr Ser Thr Gly Thr Thr Gly Pro Pro Val Thr Met Thr
178      195             200             205
179      Thr Gly Ser Leu Glu Pro Ser Ser Gly Ala Ser Gly Pro Gln Val Ser
180      210             215             220
181      Ser Val Lys Leu Ser Thr Met Met Ser Pro Thr Thr Ser Thr Asn Ala
182      225             230             235             240
183      Ser Thr Val Pro Phe Arg Asn Pro Asp Glu Asn Ser Arg Gly Met Leu
184      245             250             255
185      Pro Val Ala Val Leu Val Ala Leu Leu Ala Val Ile Val Leu Val Ala
186      260             265             270
187      Leu Leu Leu Leu Trp Arg Arg Arg Gln Lys Arg Arg Thr Gly Ala Leu
188      275             280             285
189      Val Leu Ser Arg Gly Gly Lys Arg Asn Gly Val Val Asp Ala Trp Ala
190      290             295             300
191      Gly Pro Ala Gln Val Pro Glu Glu Gly Ala Val Thr Val Thr Val Gly
192      305             310             315             320
193      Gly Ser Gly Gly Asp Lys Gly Ser Gly Phe Pro Asp Gly Glu Gly Ser
194      325             330             335

```

RAW SEQUENCE LISTING

DATE: 08/03/2005

PATENT APPLICATION: US/10/528,948

TIME: 13:57:54

Input Set : N:\Crf3\RULE60\10528948.raw

Output Set: N:\CRF4\08032005\J528948.raw

```

195      Ser Arg Arg Pro Thr Leu Thr Thr Phe Phe Gly Arg Arg Lys Ser Arg
196                340                345                350
197      Gln Gly Ser Leu Ala Met Glu Glu Leu Lys Ser Gly Ser Gly Pro Ser
198                355                360                365
199      Leu Lys Gly Glu Glu Glu Pro Leu Val Ala Ser Glu Asp Gly Ala Val
200                370                375                380
201      Asp Ala Pro Ala Pro Asp Glu Pro Glu Gly Gly Asp Gly Ala Ala Pro
202      385                390                395                400
204 <210> SEQ ID NO: 5
205 <211> LENGTH: 1924
206 <212> TYPE: DNA
207 <213> ORGANISM: Human leukosialin
208 <400> SEQUENCE: 5
209      cctctgagcc cagccctccc tagcatcacc acttccatcc cattcctcag ccaagagcca      60
210      ggaatcctga ttccagatcc cacgcttccc tgccctccctc aggtcccagc tcttgctcct      120
211      gcctgtttgc ctggaaatgg ccacgcttct ccttctcctt ggggtgctgg tggtaagccc      180
212      agacgctctg gggagcacia cagcagtgcg gacacccacc tccggagagc ctttggtctc      240
213      tactagcgag cccctgagct caaagatgta caccacttca ataacaagt accctaaggc      300
214      cgacagcact ggggaccaga cctcagccct acctccctca acttccatca atgagggatc      360
215      ccctcttttg acttccattg gtgccagcac tggttccctt ttacctgagc caacaacct      420
216      ccaggaagtt tccatcaaga tgtcatcagt gccccaggaa acccctcatg caaccagtca      480
217      tcctgctgtt ccataacag caaactctct aggatccac accgtgacag gtggaacct      540
218      aacaacgaac tctccagaaa cctccagttag gaccagtggg gccctgttta ccacggcagc      600
219      tagctctctg gagacctcca gaggcacctc tggacccctt cttacctagg caactgtctc      660
220      tctggagact tccaaaggca cctctggacc ccctgttacc atggcaactg actctctgga      720
221      gacctccact gggaccactg gacccctctg taccatgaca actggctctc tggagccctc      780
222      cagcggggcc agtggacccc aggtctctag cgtaaaacta tctacaatga tgtctccaac      840
223      gacctccacc aacgcaagca ctgtgccctt ccggaaccca gatgagaact cagcaggcat      900
224      gctgccagtg gctgtgcttg tggccctgct ggcggtcata gtccctgtgg ctctgctcct      960
225      gctgtggcgc cggcggcaga agcggcggac tggggccctc gtgctgagca gaggcggcaa      1020
226      gcgtaacggg gtggtggacg cctgggctgg gccagcccag gtccctgagg agggggccgt      1080
227      gacagtgacc gtgggagggg ccgggggcga caagggtctt gggttccccg atggggaggg      1140
228      gtctagcgtg cggcccacgc tcaccacttt ctttggcaga cggaagtctc gccagggctc      1200
229      cctggcgatg gaggagctga agtctgggtc agggcccagc ctcaaagggg aggaggagcc      1260
230      actggtggcc agtgaggatg gggctgtgga cgccccagct cctgatgagc ccgaaggggg      1320
231      agacggggct gccctttaag tgtcggtgaa tagtgaggct ggaggccgca atctcagcca      1380
232      gcctccagca ctttccctct caccatccca ctgccccctc gctcccatgt ttccaccggg      1440
233      caccctgac ctcacccgaa tctccttttt tttttcttt tgagacagag ttctgctttg      1500
234      tcgcccaggc tggagtgcaa tgcacgatct cagttcactg caacctctgc ctctaagtt      1560
235      caggcgattc tcctgcctca gcttcccgag taactgagat tacaggcacc caccaccatg      1620
236      cccagctgct tttttgtatt tttggtagag atggggtttc accatgttgg ctaggctggt      1680
237      ctcaaactcc tgacctcagg tgatctacct gcctcagcct cccaaagtgc tgagattaca      1740
238      gacatgagcc tccgcgcctt gcctcctcac ccacctcttc actctgaatc ctcatgagge      1800
239      ttctcagccc tggatttctt gctgccatcc tcacccagca cccacaacta gcgcctgggc      1860
240      agggcagggc tggcacctct caacgtctgt ggactgaatg aataaacctt cctcatccac      1920
241      ccct
242
243 <210> SEQ ID NO: 6
244 <211> LENGTH: 400
245 <212> TYPE: PRT

```

RAW SEQUENCE LISTING ERROR SUMMARY

DATE: 08/03/2005

PATENT APPLICATION: US/10/528,948

TIME: 13:57:55

Input Set : N:\Crf3\RULE60\10528948.raw

Output Set: N:\CRF4\08032005\J528948.raw

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:1; Line(s) 4

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/528,948

DATE: 08/03/2005

TIME: 13:57:55

Input Set : N:\Crf3\RULE60\10528948.raw

Output Set: N:\CRF4\08032005\J528948.raw